Solving Probation and Change-of-Major Issues in Higher Education Using Data Mining Techniques

Marie Khair, Walid Zakhem and Chady El Moucary

Notre Dame University – Louaize, North Lebanon Campus, Barsa, El Koura, North, Lebanon

***Abstract*— the rate of success in higher education is of utmost importance for both students and academic institutions. Indeed, it is a turning point in the students' life and a prerequisite for entering the professional world and/or developing a career. Additionally, higher-education institutions are competing in keeping a high retention rate and adequate transcripts with high students' performance. Admissions' entrance exams, school grades and SATs are some of the screening procedures that institutions exercise in order to monitor students through different programs of studies. Nonetheless, students are not usually fully aware of their choice when they apply for a major in terms of the requirements it shall imply, namely regarding special skills, which reveal only at advanced stages of their studies. In many cases, students fall under probation, face suspension and eventually have to change their major. The objective of this paper is to draft a clear and accurate assessment procedure that shall anticipate the aforementioned predicaments. The procedure will be developed using Educational Data Mining techniques and Neural Networks with the perspective of endowing both students and advisors the capacity of an early-stage decision-making tool based on the performance during the first year of studies and using specific courses, "*relieffed*" as indicators. Simulation results show the effectiveness of the suggested approach in terms of predicting academic issues at very early stages and offering pertinent guidelines and solutions thus, sustaining a higher retention rate in the different programs and majors.**

***Keywords*–** Component, Educational Data Mining, Neural Networks, Clustering, Relief, Probation, Change-of-Major and Students' Performance

I. INTRODUCTION

Higher Education requirements have shifted with the dawn of the current millennium and International Accreditation seems to be a mandatory path for quality assurance and for sustaining an acceptable reputation and a place on the map given the merciless competition. Crucial main strains in the specifications' long list are the rates of students' acceptance, retention, graduation and performance. These well-articulated criteria have a common denominator; appropriately guiding students through their majors of choice [1], [12].

In many cases, namely in those majors who are rated high in demands due to job markets' trends, students enroll without fathoming enough the special skills and background-capabilities these majors imply, or merely because they were not cognizant of what career they are likely to embrace in their near future. This results in students being ill-guided, if not misled. The repercussions might be disastrous especially when students reach advanced levels in their curricula and face unavoidable situations such falling under probation, getting suspended or worse, having to change their major. This directly affects the aforementioned requirements that higher-education and academic institutions are attempting to improve or avoid.

Fortunately, the Higher Education Business has finally awakened and started resorting to Educational Data Mining (EDM) approaches and techniques, which have proven to be decisive when tackling complex and vital academic issues. Indeed, colleges and universities have grown to a point where native/individual skills or simplistic methods do no more apply for the large amount of information and high level of interrelatedness such databases present to analysis. Furthermore, Institutional Research Office employing data miners and expert analysts is becoming an integral part of every institution seeking to grow and expand while sustaining good academic standards and acceptable pertinent rates.

The objective of this paper is to draft a systematic and accurate predictive algorithm that shall anticipate the aforementioned academic predicaments students might face by particularly avoiding difficult and late-stage decisions that the institution has to proceed with. The algorithm will be developed using Educational Data Mining techniques and Neural Networks. The main focus is to underline those courses, which we call indicators that mostly affect students' performance, i.e. General Point Average (GPA), in every year of the suggested program of a given major of studies. To this purpose, *Relief* will be exploited to estimate the weight of contribution of each course towards students' GPA, namely those courses that are directly related to the area of inquiry, i.e., concentration requirements. At a second stage, Neural Networks and K-Means Clustering will be employed to predict the student's yearly and overall GPA based his/her performance in such courses.

The outcome of this procedure will enable students, instructors, and advisors to benefit from the capability of underlying the courses that students have to emphasize on and dedicate extra efforts for and predict whether a student will successfully achieve his graduation. Indeed, the performance of a student in these indicators will dictate to a far extent his/her ability to pursue the program without going under probation, thus keeping a high retention rate. Additionally, students are warned in advance of these courses and thus, advisors would be able to carry on their duties in a more confident and efficient manner. Finally, the suggested technique will also estimate students' yearly GPA as well as their overall GPA upon graduation. Furthermore, this capacity

can be wisely employed in revisiting the acceptance requirements for each major.

Simulation results were carried out on a moderately large sample of students. The high percentage of accuracy and low error rates are mainly due to the adjustments and preprocessing stages that were accomplished prior to applying EDM techniques, namely Neural Networks, Relief and K-Means Clustering. These results demonstrate the effectiveness of the proposed solution as shown and discussed at a later stage.

The paper will be divided into four sections. In the first section, Educational Data Mining will be recalled and some of the techniques will be briefly presented with pertinent references. Additionally, Neural Networks and K-Means Clustering will be elaborated in the view of the application. In section III, the problem statement will be thoroughly presented and discussed. The data preparation process will be also explained with the subsequent pertaining phases. The predicting algorithm will be achieved and tested in this section as well. In section IV, simulation results will be carried out to show the efficiency and usefulness of the suggested strategy. The main objective of the paper will be highlighted in terms of early discovery of academic performance issues and appropriate measures and recommendations are offered as a tool for remedy and decision making. Finally, some conclusions and work perspectives are drawn in the light of the results.

II. BACKGROUND

A. Data Mining

The most commonly applied Data Mining (DM) tasks are regression, clustering, classification, and association-rule mining; and the most used DM techniques/methods are decision trees, neural networks, and Bayesian networks. Moreover, DM reveals completely new, hidden, and interesting information found in data [3].

Most of the DM algorithms need to be configured before they are executed. Users have to provide appropriate values for the parameters in advance in order to obtain good results/models, and therefore, the user must possess a certain amount of expertise in order to find the right settings [11], [13].

Many techniques have been created and applied for data mining and applying the best method for a particular situation is a task of its own. Many types of models have been developed, including Neural Networks (NN). Each method has its advantages and disadvantages and there is no single method that is best for all applications [17], [19].

B. Neural Networks

The name neural network was a great success of the twentieth century; it is "A network of weighted, additive values with nonlinear transfer functions" [3].

All neural networks have an input layer, hidden layer(s), and an output layer. Figure 1 is a diagram of a network with one hidden layer for a total of three layers. When there is more than one hidden layer, the output from one hidden layer is fed into the next hidden layer and separate weights are applied to the sum going into each layer [5].

A vector of predictor variable values (x1..xp) is presented to the input layer which distributes the values to each of the neurons in the hidden layer which is multiplied by a weight, and the resulting weighted values are added together producing a combined value. The outputs from the hidden layer are distributed to the output layer which is also multiplied by a weight, and the resulting weighted values are added together producing a combined value. These values are the outputs of the network [4], [5].

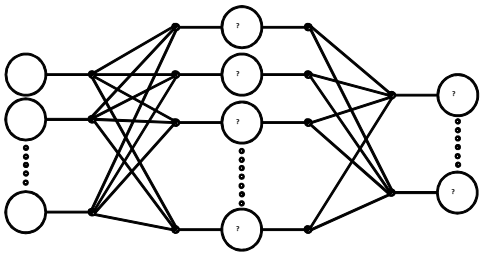

Figure 1: One Layer Neural Network

C. K-Means Clustering

Clustering is the process of grouping N patterns into different clusters based on a suitable notion of closeness or similarity among these patterns. Good clusters show closer similarity in a group and low similarity between data being fit to two different groups [11], [16], [20], [23].

D. Educational Data Mining

Educational Data Mining (EDM) is the extraction of hidden information from large sets of educational data. These sets of data comprise of the previous students, their courses taken during any one semester, one year or even two years. These data belong to real students who passed through some similar degrees, taken some similar courses, received some similar grades and whether they were able to graduate or not. Some of these students might have requested some help to improve their situations by increasing their GPA and graduate properly. Some other students did not request help, they simply did not realize that they needed assistance, they simply could not graduate. But if they knew their situation might deteriorate enough to disallow them to graduate, they would have requested help or even they were offered help by their advisors [7], [10], [18].

Predicting the student's performance is the most popular Data Mining application in higher education where universities usually try to foresee their students' performance based on their knowledge, their scores and their GPA's. Different models of Data Mining are used, one of which is Neural Networks (NN) [8], [14], [18].

The so called "machine learning" uses artificial intelligence to introduce rules that can be applied to new data. Once a model is trained, one can feed in a new student group, and the model applies the learned information to the new group to predict the likelihood of receiving High GPA. When institutions use data mining to predict which students are most at risk, institutions can prevent a student from failing before the student is even aware that he or she is at risk [5], [9].

E. Relief

The key idea of the original *Relief* algorithm is to estimate the quality of attributes according to how well their values distinguish between the instances that are near to each other it searches for its nearest neighbors: nearest hits misses and average them to update the quality estimation of all the attributes [6], [22].

Relief algorithms are general and successful attribute estimators. They are able to detect conditional dependencies between attributes and provide a unified view on the attribute estimation in regression and classification.

III. OVERVIEW OF THE PROPOSED SOLUTION

A. General Outline

The advising system is mainly based on manual study of each student by himself. The ability to create an automatic system that helps the advisor to provide a warning to each student based on his / her course load and the importance of the courses taken would facilitate the advising hassle as well as make the prediction of the drop out percentage of each student.

An engineering student at Notre Dame University is required to take several courses in a prerequisite states whereas he is required to have completed Circuits I, Circuits II, Electronics I and Electronics II in series as an example followed by several courses and at the same time he must have completed all the Mathematics, Physics and Chemistry courses just prior to pursuing the rest of the curriculum courses.

It is worthwhile noting that only major and core courses are considered during this study. It is understood that Technical Elective and General Educational Requirement courses are just as important for the GPA prediction, but since they can be taken in a merely random fashion between students, they will not be considered.

In this paper, NN will be used as it simulates how students get their GPA based on specific input (high School Grades, SAT Grades), and on intermediate layers (their semesters' grades). The system is fed with training data of graduating

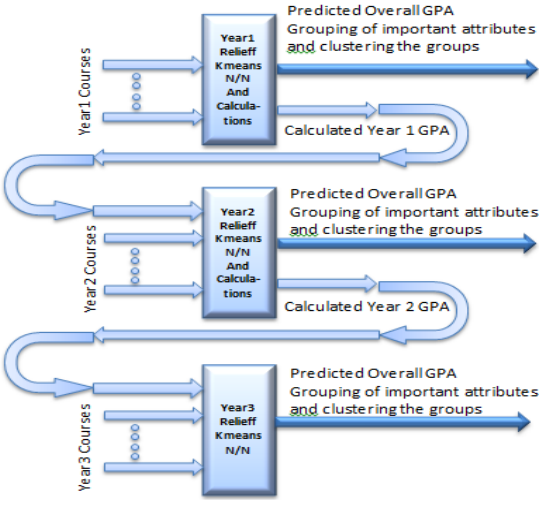

Figure 2: Block Diagram of the Proposed System

students which is 70% of the total number of records with their grades, then tested based on a testing data and validation data each is 15% of all the records.

Another important factor used in this paper is the *Relief* algorithm which is used to extract weights for the courses that the students usually take. This can be used in order to advise each student on how much effort to put on the courses taken during a semester for a better performance.

One last algorithm is the K-Means clustering, it is used to divide the courses into clusters (in this case 4). It clusters courses in such a way that students would distribute their load over courses with belonging to different clusters.

B. Block Diagram

Figure 2 shows the block diagram of the proposed system, which is having three complete layers where each layer includes NN, *Relief*, and K-Means Clustering, all using MATLAB software from Mathworks.

The system would completely be independent as for each part would be trained, tested and validated for the set of courses that belong to year1, year2 or year3. While it is also dependent as for each part would be depending on the previous calculated GPA and introduced as an attribute in addition to the already counted courses of that year toward predicting the overall GPA.

Neural Network is used to predict the overall GPA for each student based on a set of attributes that would be collected from the Registrar's Office of Notre Dame University (NDU), these attributes are divided into three years, and each year would include the courses taken as per the advised program. Should a student be enrolled in his second year, he would have already completed the courses for the first year and accumulated a Year1 GPA. Same is true for students enrolled in their third year, where they would have calculated their Year2 GPA. During the second or the third year, the NN system is to predict the overall GPA using the courses for the current year and the GPA for the previous year.

Also for the advising purpose, the student would like to know which of his courses are most important and which are not. For that reason, the courses of the current year taken by that student would be introduced to a system comprised of two blocks of which the first is to inform the student of the importance of the courses and the influencing factor of each course toward his GPA. The second block is to inform the student as well as the advisor that the courses in one group should be taken during two semesters and not all in one semester.

C. Preparing the Data

Data was gathered from 1100 students already graduated from the Electrical and Computer and Communication Engineering. Should a course be repeated, the latter grade would replace the previous one. A course can be repeated voluntarily should a student require or be required to raise his GPA, these students were excluded from this study as their final grade point average was increased voluntarily.

Furthermore, some students were at the top of their classes throughout their studies, and these students were also excluded. Preprocessing the network inputs improves the efficiency of neural network training. The data sets were cleaned and prepared and a total of 305 records were finally

ready for study excluding all the misleading records whether they are the top or the bottom of the class. The set of attributes are all courses with their letter grade as per table1 and transformed to a value that affect their final GPA by simply converting letters to values as it is stated in table 2 and the Notre Dame University Catalog [2], [24].

Table 1: Set of Attributes with Letter Grade

	Year 1									Year 2											Year 3									
	ENL 230	ENL 213	PHS 212	CHM 211	MAT 211	MAT 213	MAT 215	MAT 224	GPA Year1	PHS 213	MAT 235	MAT 335	MAT 326	CSC 212	CSC 213	CSC 312	EEN 201	EEN 202	EEN 220	EEN 210	EEN 203	EEN 221	GPA Year2	EEN 311	EEN 324	EEN 340	EEN 344	EEN 331	EEN 312	EEN 325	GPA Overall
1	A-	B+	C+	A-	A	B+	B	A	3.41	B+	A-	A	A	A	A-	B+	A	A	A-	A	A	3.79	B+	B+	B+	A-	B-	A	B+	3.58
2	B-	B	A	A-	A	A-	A	A	3.64	B	A	A	A	A	A	A	A	B+	A	A-	A-	3.81	A	A	A-	B+	C+	A-	B	3.58
3	B+	B+	C+	A-	A-	A-	A-	B	3.34	B	B	A	B+	A-	B	B+	C+	B+	B	B+	B+	A-	3.25	B+	A-	B	B	A-	B+	A-	3.49
4	C	D	C	C-	C	B	C-	C	1.93	D	B	D+	C+	B+	C-	C-	C+	B	C	C-	C+	C	2.11	C-	B+	C-	C+	C-	C-	B-	2.03
5	C-	C-	D	D	C	B	A-	B	2.14	C-	A	A-	B-	B+	B+	B+	B+	B+	B-	C-	C	B	2.87	C	C-	A	A	C	B	C	2.5
6	D	D	D+	B	C	D	C+	B+	1.86	C-	C	C-	C+	C	C	C-	C-	C-	C	C-	B	B-	2.00	C-	C+	C	C	C-	B+	B	2.11
7	B+	B	B+	D+	C-	D	B-	C	2.29	C+	D	C-	C+	B	C+	C	C+	C+	C+	B-	A	A	2.46	B	C+	C+	C+	C+	A-	B	2.6
8	C+	C	B+	B+	A	B-	C	A	2.95	B+	B	B	B-	A-	A	A	A-	B	A	A-	A-	A-	3.46	B-	B+	B	B	A	B+	A	3.29
9	B	A-	B-	A-	A	A-	B	B	3.35	C	A-	B-	B	A-	B+	C+	B+	C-	C+	C+	A	A	2.98	B	C	C	C	B	B+	B	3.06
10	B+	A-	C+	A	A	C	B	B-	3.13	D	B	C	C+	A-	B-	C+	B	C-	A	B-	A	A	2.82	B	C-	C+	C+	C+	B+	A-	2.89

Table 2: Letter Grade to Value Conversion Table

Letter	Value	Letter	Value	Letter	Value	Letter	Value
A+	= 4.0	B+	= 3.3	C+	= 2.3	D+	= 1.3
A	= 4.0	B	= 3.0	C	= 2.0	D	= 1.0
A-	= 3.7	B-	= 2.7	C-	= 1.7	F	= 0.0

D. The Proposed Algorithm

For the purpose of predicting the final GPA, a neural network was created using the "*nftool"* function of MATLAB. The NN would be trained, validated, and tested. While for the purpose of knowing the most important course during that period that is being studied, the list of inputs would be sent to the "*Relieff"* function of MATLAB that would in turn provide us with the importance of each attributes with respect to the data itself [15], [21].

That same study was done over all the years of the curriculum, the first, the second and the third, while the fourth year was not necessary to study, the student would be graduating in two semesters and the advice would not be relevant anymore.

The curriculum was divided into four parts:

- The courses taken during the first year, NN would predict the final GPA after the course is graded, "*Relieff"* would provide the most important courses taken during that period, and "kmeans" would group the importance of the courses into four groups whereas the most important course group would be advised as a set of courses that a student should handle carefully in order to raise his GPA or to keep it high enough till the following year. A GPA attribute is to be calculated from the list of grades for year1 courses.
- The courses taken during the second year as well as the GPA for the courses taken during year1 are the new attributes for the neural network data that should be studied to provide the predicted overall GPA, while "*Relieff"* would provide the most important courses during that period as well as the "kmeans" groups of the courses/attributes into four clusters of importance. Again, a new attribute, GPA for year2 is to be created for the use in the year three study.
- The courses taken during year3 as well as the GPA of year2 are the bases of the same study done for the year2.
- While the courses taken for the final are not relevant to the study, there are no expectation and no prediction for the final year courses. The system would simply add the courses and find the final GPA and know whether the student would graduate or not.

IV. RESULTS' ANALYSIS

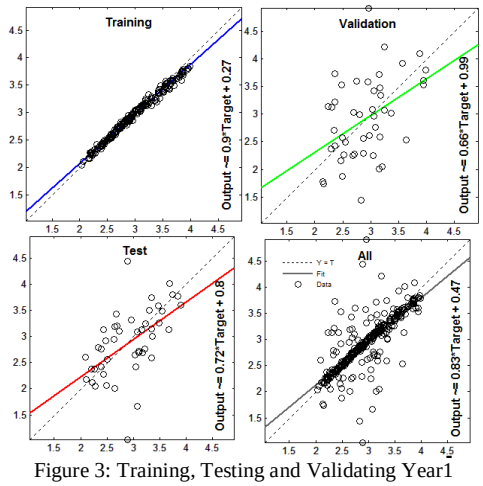

Figure 3: Training, Testing and Validating Year1

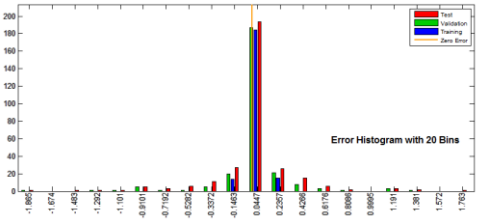

Figure 4: Error Histogram of Year1

Figure 3 and Figure 4 show the training, testing, validating and the error Histogram of the NN for the year1 where the attributes are all the student courses taken during Year 1 and after being graded. The NN would predict the overall GPA. The figures shows error of less than 17% and most errors are within 5% from the target GPA.

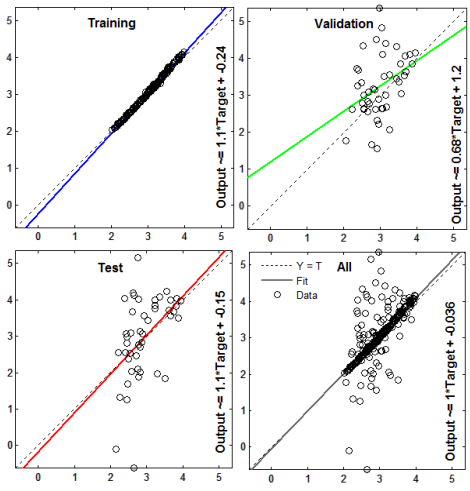

Figure 5: Training, Testing and Validating Year2

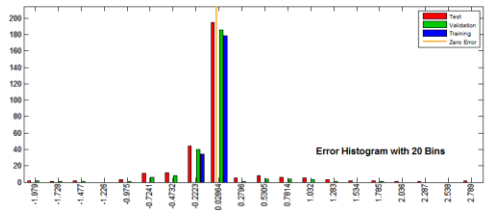

Figure 6: Error Histogram of Year2

Figure 5 shows the training, testing and validating of the NN for Year2 where the attributes are all the grades of the second year courses as well year1 GPA. The predicted attribute is the overall GPA. Figure 6 shows the error histogram where most errors are within the 2.9%

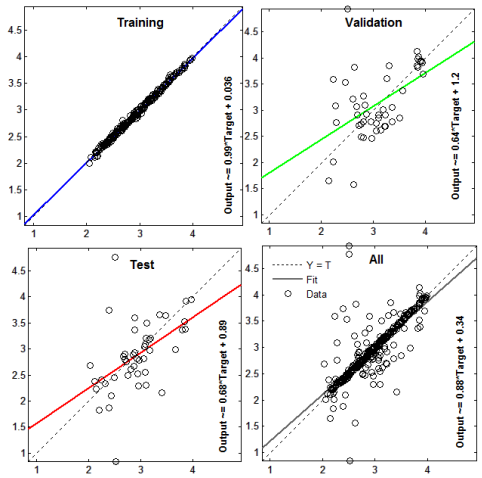

Figure 7: Training, Testing and Validating Year3

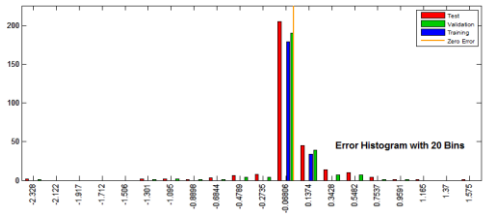

Figure 8: Error Histogram of Year3

Figure 7 shows the training, testing, and validating of the NN that has attributes the grades of the third year as well as the GPA of up to the second year courses. The predicted attribute is the overall GPA. Figure 8 shows the error histogram where most errors are within the 6.9%.

Table 3: Importance of Courses taken during Year 1

Course Number	*Course Title*	*% Weight*	*Cluster*
ENL230	English in the workplace	26.06%	1
MAT211	Discrete Mathematics	16.97%	
PHS212	Electricity and Magnetism	15.76%	
MAT224	Calculus IV	10.51%	2
ENL213	Sophomore English Rhetoric	9.90%	
MAT213	Calculus III	9.70%	
CHM211	Principles of Chemistry	7.47%	3
MAT215	Linear Algebra	3.64%	4

Table 4: Importance of Courses taken during Year 2

Course Number	*Course Title*	*% Weight*	*Cluster*
GPA Y1	GPA for Year 1	27.13%	1
PHS213	Modern Physics	15.26%	2
EEN210	Electronic Circuits I	11.42%	
MAT326	Probability & Statistics For	8.42%	3
CSC213	Program Design and Data	8.27%	
MAT335	Partial Differential Equations	6.84%	
EEN202	Circuits Analysis II	6.54%	
EEN220	Introduction to Logic Design	5.19%	
MAT235	Ordinary Differential Equations	5.19%	
CSC212	Program Design and Data	4.06%	
CSC312	Computer Architecture	2.48%	4
EEN201	Circuits Analysis I	0.34%	
EEN203	Circuits Laboratory	-0.50%	
EEN221	Logic Design Laboratory	-0.62%	

Table 5: Importance of Courses taken during Year 3

Course Number	*Course Title*	*% Weight*	*Cluster*
GPA Y2	GPA for Year 2	74.25%	1
EEN311	Electronic Circuits II	16.94%	2
EEN340	Signals and Systems	12.00%	
EEN324	Microprocessor System Design	3.18%	3
EEN331	Electromagnetics II	1.29%	
EEN344	Communication Systems I	0.21%	
EEN325	Microprocessor Laboratory	-3.41%	4
EEN312	Electronic Circuits Laboratory	-4.47%	

"Relieff" function of MATLAB was performed on the courses taken during Year1, Year2 and Year3. The courses taken during Year2 will also include the GPA of the courses taken during Year1; same is true for the Year3 whereas it would include the courses taken during that year as well as the GPA for the Year2. The lists of attributes for every year studied were clustered using "kmeans" function of MATLAB and were grouped into four clusters: "Most important, important, less important and least important" [7].

Tables 3, 4 and 5 show the importance of the courses for the Year1, Year2 and Year3. They also show clearly that the most important attribute for Year2 is the GPA for Year1 as well as the most important attribute for Year3 is GPA for the previous year.

The advantage of the proposed tool is divided into three parts, concerning the students, it helps them guess their overall GPA prior to their graduation and provides guidance, based on year by year manner, to who are expecting themselves to receive low GPA that will not allow them to graduate and put them in a probation status and further in a suspension status.

Concerning advisors, this tool helps them foresee the courses and their importance toward each and every student, the advisors can clearly and scientifically advise the student to make more effort on studying for that specific and important course and not to ignore the other courses as their yearly GPA will definitely affect their final and overall GPA.

Concerning universities, this tool created a system that foresees the students graduation GPA starting from year 1 courses, not only this but the prediction will also evolve as student will know ahead of time whether they should request help in their studies. Also the curriculum committees will benefit from the system by knowing the importance of each

course and which to exclude from one curriculum to add a fresh new topic. The faculties would also benefit as they would know which student should be guided prior to his drop-out or even before he changes major not knowing that only more effort should be exerted in order to graduate.

V. CONCLUSION

This paper presented an efficient tool that endows students and advisors as well as universities with the capacity of anticipating and thus, avoiding serious academic predicaments, namely probation and change-of-major. The tool was developed by applying Neural Networks and other Educational Data Mining tasks. Additionally, *Relief* was exercised in order to segregate significant indicators that shall guide the tool in the prediction stages.

The purpose of the study was twofold. First, the developed algorithm underlines the courses that students, instructors and advisors are to consecrate extra attention for their high correlation with students' GPA. Second, prediction of students' ongoing overall performance is predicted yearly with an accurate figure of the GPA upon graduation. These outcomes enable advisors to significantly forestall critical situations and thus, carry out preventative measures and advice.

This study provides a scientific proof that the courses taken during the first year are just as important as the last year's courses, since the GPA of the first year is the most important factor of the second year courses prediction to the final GPA. And also the second year GPA with the courses taken in the third year is the most important factor of the prediction of the GPA. There are some courses that are next to the importance of the GPA that the student must study harder and receive a high grade in order not affect his final GPA.

REFERENCES

[1] C. El Moucary, M. Khair and W. Zakhem, "Improving students performance using data clustering and neural networks in foreign-language based higher education", The Research Bulletin of JORDAN ACM, vol. 2, pp. 27-34, September 2011.

[2] C. El Moucary, "Data mining for engineering schools predicting students' performance and enrollment in masters programs," International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10, pp. 1-9, November 2011.

[3] C. Romero, S. Ventura, "Educational data mining: a review of the sate-of-the-art",. IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 40, No. 6, pp. 601-618, Novemver 2010.

[4] J. Luan, "Data mining applications in higher education", SPSS.

[5] DTREG Multilayer perceptron neural networks, http://www.dtreg.com /mlfn.htm.

[6] Y. Sun, D. Wu, "A RELIEF based feature extraction algorithm", Society for Industrial and Applied Mathematics

[7] F. Araque, C. Roldán and A. Salguero, "Factors influencing university drop out rates", ScienceDirect Computers & Education

[8] G.M. Kanakana1* and A.O. Olanrewaju2 , "Predicting student performance in engineering education using an artificial neural network at tshwane university of technology", Stellenbosch, South Africa, ISEM 2011 Proceedings, September 2011,

[9] M. Xenos, C. Pierrakeas and P. Pintelas "A survey on student dropout rates and dropout causes concerning the students in the Course of informatics of the Hellenic Open University", Computers & Education Vol. 39, pp. 361–377, 2002

[10] S. Ayesha , T. Mustafa , A. R. Sattar and M. Inayat Khan , "Data mining model for higher education system", European Journal of Scientific Research,Vol.43 No.1, pp.24-29, 2010

[11] Ryan S.J.d. Baker, "Data mining for education", Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

[12] N. Romashkin, D. Ignatov and E. Kolotova , "How university entrants are choosing their department? Mining of university admission process with FCA taxonomies" National Research University – Higher School of Economics, Moscow, Russia

[13] B. Zhao, J. T. Kwok and C. Zhang, "Multiple kernel clustering" Society for Industrial and Applied Mathematics.

[14] J. Zimmermann, k.H. Brodersen,J.-P. Pellet, E. August, and JM Buhmann, "Predicting graduate-level performance from undergraduate achievements" Department of Computer Science ETH Zurich

[15] Mathworks, "Statistics toolbox user's guide"

[16] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An Efficient k-Means clustering algorithm: analysis and implementation", IEEE transactions on pattern analysis and machine intelligence, Vol. 24, No. 7, July 2002

[17] Kurt Thearling, "An Introduction to Data Mining"

[18] J. Ranjan [1], R. Ranjan, "Application of data mining techniques in higher education in india", Journal of Knowledge Management Practice, Vol. 11, Special Issue 1, January 2010

[19] Nong Ye, "The Handbook Of Data Mining"

[20] L.S.Affendey, N. Mustapha, Nasir Sulaiman and Z. Muda, " Ranking of influencing Factors in predicting Students' Academic Performance" Information Technology Journal Vol. 9, No. 4 pp. 832-837, 2010

[21] Mathworks: "Neural network toolbox user's guide"

[22] I. Kononenko, E. Simec, M. R.-˘Sikonja, "Overcoming the myopia of inductive learning algorithms with RELIEFF", Kluwer Academic Publishers, Boston

[23] S. Trivedi, Z. A. Pardos, G. N. Sárközy and N. T. Heffernan, "Spectral Clustering in Educational Data Mining", Department of Computer Science, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA – 01609. United States

[24] Notre Dame University Catalog , http://www.ndu.edu.lb.